

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 February 2002 (21.02.2002)

PCT

(10) International Publication Number
WO 02/14860 A1

(51) International Patent Classification⁷: G01N 33/50, (74) Agent: DAVIES, Jonathan, Mark; Reddie & Grose, 16 C12Q 1/68, C12N 15/10 Theobalds Road, London WC1X 8PL (GB).

(21) International Application Number: PCT/GB01/03657

(22) International Filing Date: 15 August 2001 (15.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

0020016.2	15 August 2000 (15.08.2000)	GB
0029309.2	1 December 2000 (01.12.2000)	GB
0106610.9	16 March 2001 (16.03.2001)	GB
0113883.3	7 June 2001 (07.06.2001)	GB
0117232.9	14 July 2001 (14.07.2001)	GB

(71) Applicant (for all designated States except US): DISCERNA LIMITED [GB/GB]; Babraham Hall, Babraham, Cambridge CB2 4AT (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): HE, Mingyue [CN/GB]; 27 The Close, Babraham, Cambridge CB2 4AQ (GB). TAUSSIG, Michael, John [GB/GB]; Exmoor Cottage, Hildersham, Cambridge CB1 6BU (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SI, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 02/14860 A1

(54) Title: FUNCTIONAL PROTEIN ARRAYS

(57) Abstract: The present invention relates to a method for producing a protein array starting from DNA (or mRNA) in which a number of native, functional proteins, domains or peptides are produced in parallel by *in vitro* synthesis using a cell free system for transcription and translation. The products are immobilised in a gridded format on a surface, using an isolation sequence tag incorporated into the proteins.

FUNCTIONAL PROTEIN ARRAYS**BACKGROUND**

An array is a precisely ordered arrangement of elements, allowing them to be displayed and examined in parallel (1). It usually comprises a set of individual species of molecules or particles arranged in a regular grid format; the array can be used to detect interactions, based on recognition or selection, with a second set of molecules or particles applied to it. Arrays possess advantages for the handling and investigation of multiple samples. They provide a fixed location for each element such that those scoring positive in an assay are immediately identified; they have the capacity to be comprehensive and of high density; they can be made and screened by high throughput robotic procedures using small volumes of reagents; and they allow the comparison of each assay value with the results of many identical assays. The array format is well established for global analysis of nucleic acids, and oligonucleotide and cDNA arrays (DNA chips) are used for gene expression analysis. In a familiar format, large numbers (e.g. thousands) of DNA hybridisation probes are attached in an ordered pattern to a surface such as nylon, glass or silicon and hybridised to fluorescently labelled whole cell mRNA or cDNA; the quantitative signals on each array element are measured in parallel by means of a reader device.

The array approach may also be adapted for display of peptides and proteins; the elements displayed may be a set of related proteins or peptides, or the entire protein complement of an organism. Protein array technology allows high throughput screening for gene expression and molecular interactions. It is possible to use protein arrays to examine in parallel the functions of thousands of proteins previously known only by their DNA sequence. For functional information to be obtained, the arrayed proteins must be in native form. However, some preparative methods cause protein denaturation, as may occur during extraction or release of recombinant proteins from bacteria, and the use of arrays from such starting material is therefore limited to applications determined only by the primary sequence of the protein rather than tertiary structure. In order to develop high throughput approaches to global protein analysis which can yield functional information, methods for producing arrays in which proteins retain their functions are required.

Arrays of immobilised proteins can be used to demonstrate a binding reaction, as where the array is exposed to an entity such as an antibody or ligand, which may be directly or indirectly labelled, and binding demonstrated by localisation of the label to a particular segment of the array. Mass spectrometry may also be used to identify binding interactions on the array. Alternatively, the arrayed proteins may be in solution and used to study biochemical function. Potential uses of protein arrays which have been discussed in the literature (1-12) include identification of antibodies and analysis of antibody specificity, measurement of global protein expression, identification of ligand-receptor interactions and protein-protein interactions, and screening and selecting proteins or ligands from libraries. (i) *Expression profiling*. One type of protein array which has been proposed is based on immobilisation of antibodies at a surface (an antibody array). In principle, the reaction of an antibody array with cellular proteins can provide a global quantitative readout of all the proteins expressed at any particular time (proteome analysis). In one version, for differential display, the array is probed with fluorescently labelled proteins from two different cell states; cell lysates are labelled by different fluorophores and mixed such that the colour acts as a readout for the change in abundance. (ii) *Antibody detection*. A second application is the detection of antibodies against cellular proteins, where either or both partners are unknown. Thus an array of cellular proteins can be used to select antibodies from libraries of soluble antibodies or from phage-display or ribosome-display libraries. The antigen array can also be used to analyse antibodies in small amounts of patient sera, as during infections or in autoimmune conditions. (iii) *Ligand screening*. An array of potential target proteins, such as receptors, can be used as a screen for selection of ligands which may be possible drug candidates, including small molecules, peptides, aptamers, nucleic acids or synthetic scaffolds. (iv) *Detection of protein-protein interactions* A further use for protein arrays is in the detection of protein-protein interactions. Each protein in the genome may interact with a number of partners, so for the approximately 100,000 proteins encoded in the human genome there may exist millions of interactions. Such interactions are often measured by yeast two-hybrid (cell-based) methods but these may fail to measure interactions involving secreted proteins, proteins with disulphide bridges and membrane bound proteins such as receptors. An array method would be highly desirable in these cases and may reveal interactions which are not detected by the cellular methods.

Literature descriptions of preparation of protein arrays

The arrays described to date are composed either of purified proteins or proteins expressed in living cells or viruses. Early examples were peptide arrays, in which peptides were chemically synthesised on a solid support and used to identify epitopes recognised by antibodies (2). Peptide arrays can be chemically synthesised up to a length of about 30 amino acids, but are unable to produce full length folded proteins.

Clearly, protein arrays can be made by chemical or noncovalent attachment of preformed proteins onto suitable surfaces, such as treated glass slides or absorptive membranes such as nitrocellulose or PVDF. For high throughput studies, such as proteomics or library screening, this requires methods for the preparation, purification and immobilisation in parallel of large numbers of proteins. Methods for production of recombinant proteins from bacteria for assay in an arrayed format have been described (3,6-10). Proteins can be expressed as constructs fused to an affinity tag (e.g. hexahistidine) or glutathione S-transferase (GST), recovered by cell lysis and used either as crude lysates or after affinity purification (e.g. on Ni-NTA metal affinity columns). However, the production of recombinant proteins in bacterial systems can be problematic due to aggregation, insoluble inclusion bodies and/or degradation of the product, while eukaryotic systems suffer from lower yields and high demands on sterility or time consuming cloning procedures (e.g. Baculovirus). Where denaturants are used in the extraction the protein will often be rendered nonfunctional. Once the proteins have been isolated, various technical formats, substrates, production methods and detection systems are available (reviewed in 8).

Martzen *et al.* (3) purified most of the soluble yeast proteins from *Saccharomyces cerevisiae* by glutathione agarose affinity chromatography from 6144 yeast strains each of which contained a plasmid with a different yeast ORF (open reading frame) fused to GST. Proteins were assayed in solution for a particular enzymatic activity. Since the proteins were purified in native form, this constituted a functional protein array, although the proteins were not immobilised on a surface. Yeast cells have been used to create a 'living' recombinant protein array, containing about 6000 colonies, each of which expresses a different ORF Gal4-fusion protein (4). This is basically a cellular, yeast two hybrid screen performed in a 96 well plate format.

Arrays can be prepared by inducing the simultaneous expression of large numbers of cDNA clones in an appropriate vector system and high speed arraying of protein products. Bussow *et al.* (6) arrayed proteins expressed from cDNA clones of a human fetal brain cDNA expression library hEx1 cloned in *Escherichia coli*. The his6 (hexahistidine)-tagged proteins were induced from individual colonies grown in 384-well microtiter plates and gridded onto high density PVDF filter membranes prior to expression induction with IPTG. Release of the proteins from the bacterial cytoplasm created an array of proteins immobilised on the filters. Two example proteins were identified on the filters using antibodies. While this method allows the operator to screen expression libraries, the extraction procedure used 0.5M NaOH, during which process proteins were denatured and therefore rendered nonfunctional. In another report (below), proteins were extracted and solubilised using 6M guanidinium HCl, which is also a denaturant. Other drawbacks of this procedure as a means of producing an array are that clones must be extensively screened for in frame expression and that cDNA libraries contain many clones which lack the 5' end (N-terminal), may have multiple copies of some genes and poor representation of others.

Lueking *et al.* (9) gridded purified protein solutions from the hEx1 library onto PVDF filters, in an extension of classical dot-blotting methodology. For high throughput, small-scale protein expression, clones of the hEx1 library were grown in 96-well microtiter plates and induced with IPTG; cells were lysed with 6M guanidinium HCl and supernatants filtered through a 96 well filter plate onto a PVDF membrane. For larger scale production of purified proteins, peptide- and his6-tagged proteins were expressed from *E. coli* and isolated with Ni-NTA agarose. Results of high throughput screening showed quite a number of false positives, i.e. proteins detected by anti-tag which were in fact out of frame with the tag, and antibody specificity screening often showed unexpected crossreactions, mostly with ribosomal proteins, for no apparent reason. The use of guanidinium HCl denatures the proteins and may cause aberrant results.

Holt *et al.* (7) screened the hEx1 library to identify specific antibodies, reactive with denatured proteins, using 12 well-expressed antibody fragments of previously unknown specificity. Four specific interactions were identified

In another example of dot blotting, Ge described an array system for detection of protein

interactions with other proteins, DNA, RNA and small ligands (12). In this case, 48 highly purified, native proteins were arrayed on a nitrocellulose membrane, by spotting using a 96-well dot blot apparatus. The proteins were overexpressed in bacteria or baculovirus and purified to homogeneity. The dot blot array was reacted with a number of different radiolabelled probes (protein, DNA, RNA, ligand), followed by autoradiography and densitometry, and showed to behave in a functional manner, i.e. probes interacted with partner molecules with the expected specificity.

Afanassiev *et al.* (5) describe a method for making protein arrays using chemical coupling of proteins to an agarose film on microscope slides; the agarose is activated by sodium periodate to reveal aldehyde groups which bind amino groups on the protein. Varying amounts of an antigen (BAD) or an anti-BAD antibody (6A11) were immobilised and binding of the partner molecule detected by a fluorescent second reagent.

References

1. Emili A.Q. and Cagney G. (2000) Large-scale functional analysis using peptide or protein arrays. *Nature Biotechnology* 18:393-397.
2. Geysen H.M., Meloen, R.H. and Barteling S.J. (1984) Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc. Natl Acad. Sci. USA* 81:3998-4002.
3. Martzen M.R., McCraith S.M., Spinelli S.L., Torres F.M., Fields S., Grayhack E.J. and Phizicky E.M. (1992) A biochemical genomics approach for identifying genes by the activity of their products. *Science Nov 5; 286(5442):1153-5.*
4. Uetz P. *et al.* (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403, 623-627.
5. Afanassiev V., Hanemann V. and Wolf, S. (2000) Preparation of DNA and protein microarrays on glass slides coated with an agarose film. *Nucleic Acids Research* 28:E66.
6. Bussow K., Cahill D., *et al.* (1998) A method for global protein expression and antibody screening on high-density filters of an arrayed cDNA library. *Nucleic Acids Res.* 26:5007-5008.
7. Holt L.J., Bussow K., Walter G and Tomlinson I.M. (2000) By-passing selection: direct screening for antibody-antigen interactions using protein arrays. *Nucleic Acids Research* 28:e72.

8. Walter G., Bussow K., Cahill D., Lueking A. and Lehrach H. (2000) Protein arrays for gene expression and molecular interaction screening. *Current Opinion in Microbiology* 3:298-302.
9. Lueking A., Horn, M., Eickhoff, H., Bussow, K., Lehrach H. And Walter G. (1999) Protein Microarrays for gene Expression and Antibody Screening. *Anal. Biochem.* 270:103-111.
10. Bussow K., Nordhoff E., Lubbert C., Lehrach H and Walter G. (2000) A human cDNA library for high-throughput protein expression screening. *Genomics* 65:1-8.
11. Pandey A. and Mann M. (2000) Proteomics to study genes and genomes. *Nature* 405:837-846.
12. Ge H. (2000) UPA, a universal protein array system for quantitative detection of protein-protein, protein-DNA, protein-RNA and protein-ligand interactions. *Nucleic Acids Research* 28, e3.

DESCRIPTION OF THE INVENTION

We describe a method for producing a protein array starting from DNA in which a number of native, functional proteins or domains are produced in parallel by *in vitro* synthesis using a cell free system for transcription and translation, followed by immobilisation of the products in a gridded format on a surface, using an isolation sequence tag incorporated into the proteins. In one embodiment, the array is formed *in situ* in a single step by performing protein expression in wells, on surfaces, or in the presence of beads, any of which are coated with immobilising molecules (affinity ligands) such as metal ions or antibodies.

Starting material can include genomic DNA, mRNA, cloned DNA fragments, or cDNA libraries, etc. The input DNA constructs for *in vitro* transcription/translation may be obtained by PCR (polymerase chain reaction) or RT (reverse transcription)-PCR amplification, using primers designed on any known DNA sequences, such as those from databases and genome projects. Alternatively, cloned plasmid DNA may be used. The cell free systems for protein synthesis are those such as rabbit reticulocyte, wheat germ, yeast or bacterial extracts. Numbers of individual native proteins or domains may be produced in parallel directly from the PCR DNA constructs. In one embodiment the reaction is carried out in the wells of multiwell plates. In another embodiment, proteins may be synthesised directly onto surfaces, such as glass, membranes or beads, for example by carrying out the reaction in microdroplets. The construct DNA may be added to the reaction volume or droplet or pre-immobilised to the surface. In the latter case, a DNA array is effectively created first and then used to create the protein array. The proteins can be adapted for rapid isolation, immobilisation or identification by inclusion of sequences such as hexahistidine or other peptide tags. If the wells or other surfaces in, or on, which the proteins are produced are precoated with an immobilising reagent such as nickel ions or anti-tag antibodies, the array will be formed as the proteins are produced *in situ* and reagent molecules can be removed by washing. The reaction may alternatively be performed in the presence of beads coated with an immobilising reagent, the beads being subsequently distributed in a gridded array format. Alternatively, the proteins can be transferred (e.g. by gridding robot) to secondary surfaces, such as plastic, glass, agarose, beads, nitrocellulose or other membranes.

Target molecules, such as labelled ligands, proteins or nucleic acids, may then be exposed to the array and binding to individual array locations detected by means of enzyme-coupled reaction, fluorescence, autoradiography or mass spectrometry. The arrays can thereby be used for screening of antibodies, ligands, protein interactions, etc. In some cases the arrayed proteins may be used in solution in order to perform biochemical studies. Retaining the gridded format, the solutions may also be transferred to filters or plates precoated with target molecules such as antigens, and binding detected by labelled secondary reagents.

The arrays can also be linked to library display systems. Thus, target molecules may be those of display libraries, such as phage or ribosome display libraries, in which the individual proteins are linked to encoding DNA or mRNA. After binding to the array, interacting molecules are identified by amplification and identification of the linked DNA or mRNA, for example by cloning phage or by PCR, RT-PCR, hybridisation or other methods.

By utilising *in vitro* synthesis, the above methodology provides a rapid means of obtaining functional protein arrays directly from DNA, including proteins or domains known only from DNA sequence. It can replace cloning, *in vivo* expression systems and purification procedures. It avoids the problems of inclusion bodies, aggregation and degradation often encountered with bacterial expression. Since PCR and *in vitro* translation can be carried out using many samples simultaneously in parallel, such arrays will provide a high throughput capacity for analysis of protein expression, functional activity and interactions, making use in particular of the genetic information from genome projects.

DETAILS OF METHOD

DNA constructs

The constructs for *in vitro* transcription/translation (Figure 1) include: upstream T7 promoter and protein expression signals designed to allow genes to be expressed in either eukaryotic systems (Figure 2a) or in both prokaryotic and eukaryotic systems (Figure 2b); upstream or downstream sequences encoding a flexible linker (19 residues) and a tag sequence such as (his)6 [SEQ ID No. 1] or (his)6-SRAWRHPQFGG-(his)6 [SEQ ID No. 2] for protein immobilisation by metal affinity binding (Figure 3) and/or a further peptide tag sequence for recognition by antibody or ligand; and a downstream translation stop codon. To create such constructs, the gene of interest is amplified separately by PCR or RT-PCR using specific primers derived from known DNA sequences, and the upstream and downstream elements are incorporated into the construct by PCR assembly.

For covalent immobilisation of constructs to a surface or bead, a DNA fragment encoding a transcription termination region is also included in the construct (Figure 4). In addition, a chemical group for covalent linkage, such as a terminal amino group, can be introduced via modification of the 3' primer used in PCR. The surface or bead is prepared for attachment by appropriate chemistry, such as DNA Immobiliser™ anthraquinone photo-coupling (Exiqon).

Design of PCR construct for *in vitro* expression of tagged single chain antibodies

Figure 10 outlines a general PCR strategy for construction of DNA suitable for *in vitro* protein synthesis for PISA, using single chain antibody V_H/K fragments as example. The construct contains a T7 promoter for transcription by T7 RNA polymerase and Kozak sequence for translation initiation in cell free eukaryotic systems. To increase the efficiency of protein immobilisation on a Ni-NTA coated surface and allow the re-use of the protein arrays, a double (His)₆-tag domain was designed. A flexible 19 residue linker (Robinson and Sauer, 1998, *Proc. Natl Acad. Sci. USA* 95: 5929-5934) is placed between the protein to be arrayed and the His-tag domain, in order to reduce any possible interference of the tag sequence on folding of the

-10-

attached protein. A poly(A)₂₈ tail and a transcription terminator are incorporated at the 3' end of the DNA to increase transcription efficiency. To ensure translation termination and release of the nascent polypeptide from the ribosome complex, two stop codons (TAATAA [SEQ ID No. 3]) are included following the double (His)₆-tag sequence.

A construct in which V_H/K fragments were linked to the double (His)₆-tag domain (V_H/K-His) was produced by PCR with the following primers.

Primers for PCR generation of V_H/K-linker fragments

T7 Ab/back: 5'-GCA GC T AAT ACG ACT CAC TAT AGG AAC AGA CCA CC **ATG** (C/G) AG GT(G/C) CA(G/C) CTC GAG (C/G)AG TCT GG-3' [SEQ ID No. 4]. This primer provides the T7 promoter and Kozak signal (underlined) and a degenerate sequence complementary to the 5' region of the antibody heavy chain (italics). The initiation codon is indicated in bold.

Ab-linker/for: 5'-GCC ACC GCC TCT AGA GCG GCT CAG CGT CAG GGT GCT GCT-3' [SEQ ID No. 5]. This provides a sequence complementary to the 3' region of the human κ constant domain (Cκ) and a sequence (underlined) overlapping a *linker-tag/back* primer (see below) used for generation of the double (His)₆-tag domain.

Primers for PCR generation of His-tag domain

Linker-tag/back: 5'-GC TCT AGA GGC GGT TCT GGT GGC GGT TCT GGC GGT GGC ACC GGT GGC GGT TCT GGC GGT GGC AAA CGG GCT GAT GCT GCA [SEQ ID No. 6]. This provides a sequence (underlined) overlapping the *Ab-linker/for* primer used for V_H/K-linker construction (above) and the linker sequence for PCR generation of the double (His)₆-tag domain (below).

His-tag/for : 5'-TCC GGA TAT AGT TCC TCC-3' [SEQ ID No. 7].

Plasmid PTA-His encoding a double (His)₆-tag domain

The plasmid PTA-His contains a fragment encoding (in order) a flexible linker and a double (His)₆-tag, followed by two stop codons, a polyA tail and a transcription termination region. The sequence of this fragment was: GC TCT AGA ggc ggt ggc tct ggt ggc ggt tct ggc ggt ggc acc ggt ggc ggt tct ggc ggt ggc AAA CGG GCT GAT GCT GCA **CAT CAC CAT CAC CAT CAC TCT**
AGA GCT TGG CGT CAC CCG CAG TTC GGT CAC CAC CAC CAC CAC TAA
TAA (A)₂₈ CCG CTG AGC AAT AAC TAG CAT AAC CCC TTG GGG CCT CTA AAC GGG
TCT TGA GGG GTT TTT TG CTG AAA GGA GGA ACT ATA TCC GGA [SEQ ID No. 8]. Lower case sequence: linker encoding a 19 amino acid sequence; underlined sequence: double (His)₆ tag; bold: stop codons; (A)₂₈ : poly A region which contains 28x A. Italics, underlined sequence: transcriptional terminator.

Construction of PCR fragments

In general, standard PCR was carried out for 30 cycles to obtain the V_H/K-linker fragment and double (His)₆-tag domain in separate reactions using Taq polymerase (Qiagen, UK) according to the manufacturer's instructions. The resulting fragments were analysed and eluted from a 1% agarose gel using a gel extraction kit (Qiagen, UK). For assembly, equal amounts (total 10-50ng) of V_H/K-linker and double (His)₆-tag domain were mixed and added to a PCR mixture containing 2.5μl 10x PCR buffer (supplied with Taq DNA polymerase), 1μl 2.5mM dNTPs, 1U Taq DNA polymerase and H₂O to a final volume of 25μl. After thermal cycling for 8 cycles (94°C for 30 sec, 54°C for 1 min and 72°C for 1 min), 2μl of the mixture was subjected to a second PCR in 50μl for 30 cycles using primers T7Ab/back and His-tag/for to generate V_H/K-His (see Figure 10).

In one embodiment the gridded format for transcription and translation comprises microwells of appropriate plates (e.g. 96, 384 or 1536 well polystyrene plates). Individual DNA constructs (1 μ g) are dispensed into the wells, each containing a small volume (e.g. 1-50 μ l) of cell free coupled transcription/translation systems such as the rabbit reticulocyte TNT T7 Quick for PCR DNA system (Promega), the TNT coupled wheat germ extract system (Promega) or E. coli S30 extract systems, together with methionine (0.02 mM). To label the protein, 35 S methionine or other labelled amino acids may be included. Plates are incubated for 1 hour at 30 °C. Where the proteins contain immobilisation sequences, such as single or double hexahistidine tags or specific peptide tag sequence, they may be bound either by metal affinity on surfaces or beads coated with nickel or by antibody directed against the tag sequence. Therefore, if the wells in which the translation reaction occurs are precoated with Ni-NTA (nitrilotriacetic acid), such as HisSorb plates and strips (Qiagen), or anti-tag antibody, the proteins will be bound to the surface immediately after production to generate the array *in situ*. Similarly in the presence of beads coated with Ni-NTA the protein will become bound to the surface of the bead. The TNT reaction may also be carried out in noncoated wells and the translation mixture transferred, by manual or automated procedure, to another immobilising surface, such as glass coated with Ni-NTA, beads coated with Ni-NTA, or nitrocellulose or PVDF filter membranes. The unbound material is washed away leaving bound protein.

In alternative embodiments, the transcription/translation reactions are carried out on other surfaces, for example in droplets distributed on the surface of glass, membranes or agarose. The droplets may be oil-covered to prevent evaporation.

In embodiments where protein synthesis is carried out on immobilised DNA, the DNA constructs are first attached covalently or noncovalently to the surface, which may be a polystyrene microwell, modified glass, membrane, beads, agarose, or other surface, thus forming an immobilised DNA array on which protein synthesis can occur. The surface may also be coupled to immobilising reagents such as Ni-NTA or anti-tag antibody, in order to immobilise the protein as it is produced. The protein synthesis reaction is then performed *in situ*, for example in microwells or in microdroplets of the cell free system placed onto the DNA locations on glass,

-13-

etc. Thus, in this embodiment, an immobilised DNA array is converted into a protein array via *in vitro* transcription/translation.

Array quality control

The presence of translated protein can be demonstrated using incorporated radiolabel or by antibody against a defined tag sequence shared by all the constructs. In this way the content of different wells or array locations can be normalised. Functionality may be demonstrated by specific ligand binding or enzyme activity, as appropriate to the protein.

EXAMPLES**Example 1. Creation of a functional protein array element by simultaneous expression and immobilisation of an antibody fragment *in situ*.**

To demonstrate the creation of a functional protein array element *in situ*, a construct encoding a single-chain human anti-progesterone antibody V_H/K fragment (P5-17) was used (He *et al.*, 1999, *J. Immunol. Methods*, 231:105). To prepare the construct for the protein array, the T7 promoter, expression signal and the double hexahistidine tag were incorporated into this fragment by PCR assembly. The PCR fragment (0.5 µgm) was mixed with 20µl of coupled transcription/translation 'TNT for PCR DNA system' (Promega), total volume 25µl, and the mixture was incubated at 30°C for 1 hour on Ni-NTA coated wells of an 8-well HisSorb strip (Qiagen). As control, a non-antibody PCR fragment was used in the same transcription/translation and incubation conditions. After incubation, the strips were washed three times with PBS (phosphate buffered saline), 0.05% Tween. To demonstrate that the antibody fragment was expressed and immobilised on the surface of the wells, HRP (horseradish peroxidase)-anti-human κ antibody was applied. To test the binding activity of the immobilised P5-17, the wells were incubated with biotinylated progesterone-BSA (P-BSA) for 1 hour followed by detection with HRP-streptavidin for another hour. The HRP activity was developed and the colour read by ELISA reader at 450nm.

The results of this two component array experiment show that the antibody fragment was successfully expressed *in vitro*, became bound *in situ* to the surface of the well (positive anti-κ binding) and was functional by the criterion of specific antigen binding (Figure 5). This result demonstrates that an individual protein array location can be constructed in a microwell by *in vitro* transcription and translation of PCR DNA and simultaneous immobilisation of the product to the surface of the well, with retained function of the protein.

Example 2. Creation of a functional protein array element by expression *in vitro* followed by immobilisation of an antibody fragment to a separate surface.

The P5-17 construct was transcribed and translated in microwells which were not coated with immobilising ligand, using the above conditions. After 1 hour, the well contents (translation mixture) were diluted 4-fold with PBS and 50 µl was transferred to Ni-NTAwells of HisSorb strips. After 1 hour, the wells were washed and exposed to biotinylated progesterone-BSA, BSA or CEA, followed after 1 hour by HRP-streptavidin. Figure 6 shows that the P5-17 protein was successfully produced and immobilised in the separate well, demonstrated by positive anti- κ binding, and that it was specific for progesterone-BSA. Thus a protein array element can be produced by *in vitro* transcription and translation of PCR-produced DNA and transfer of the product to the surface of an immobilising well, with retained function of the protein.

Example 3. Demonstration that immobilisation requires the hexahistidine tag.

In order to rule out nonspecific protein localisation, a mouse antibody V_H/K fragment of the anti-progesterone DB3 was produced as above using two different constructs. In one, the single hexahistidine tag sequence was included whereas in the other the sequence was deleted. Protein synthesis was carried out in HisSorb wells. Only the (His)6-containing construct was detected after translation by HRP-linked anti-mouse κ , implying that the immobilisation *in situ* requires the hexahistidine sequence.

Example 4. Functional analysis of a human single chain anti-progesterone V_H/K (P5-17) by *in situ* protein array.

A DNA construct encoding fragment P5-17 was generated by PCR. The construct contains a T7 promoter and kozak sequence for *in vitro* protein synthesis, a double His-tag for protein immobilization and a polyA tail and transcription termination region for efficient protein production. Protein expression was carried out by adding the PCR construct to Promega 'TNT Quick for PCR' kit. The mixture was incubated on individual wells of a Ni⁺-NTA coated plate (Qiagen) so that protein generation and *in situ* immobilisation proceeded simultaneously. Each well contained 25 µl of the TNT translation mixture and the mixture was incubated at 30°C for 2-3hrs with shaking. After washing with PBS-Tween (3 times), wells were treated with either biotinylated progesterone-BSA (P-BSA) or HRP-linked sheep anti-human- κ . For detection of

biotinylated antigen, HRP-linked streptavidin was used. As controls, biotinylated BSA (BSA) and anti-mouse -κ (mouse) were used. Figure 6 shows while biotinylated BSA produced no binding, the biotinylated progesterone-BSA was strongly bound (yellow colour). Similarly, the presence of P5-17 was only detected by anti-human-κ (blue) and not by anti-mouse-κ. This experiment thus demonstrates the functional expression and immobilisation of P5-17 fragment in the 'protein *in situ* array' format (PISA).

Example 5. Quantitative estimation of the expression and *in situ* immobilisation of an antibody fragment by western blotting.

The antibody V_H/K fragment produced as above was analysed by SDS-PAGE either before or after immobilisation on the microwell plate. To estimate the amount of fragment made, standard amounts of VH/K purified from *E. coli* were run alongside on the gel. Figure 7 indicate that about 150 ng V_H/K could be generated in 25 μl TNT mixture. After *in situ* immobilisation, about 50% of the total V_H/K remained in the supernatant and 30% was eluted from the plate in this case, indicating a binding of about 50ng.

Example 6. Screening of cloned antibody V_H/K fragments from a library by a functional protein *in situ* array (PISA).

The array was established by the protein *in situ* array procedure using PCR products from individual *E. coli* clones carrying DNA encoding human anti-progesterone V_H/K fragments. The clones were obtained by *E. coli* transformation with a transgenic mouse V_H/K library either before or after progesterone-BSA selection (He *et al.* 1999, J. Immunological Methods 231: 105). In Figure 8, the array elements were displayed in duplicate, with pre-selection clones numbered 1-5 (top) and post selection clones labelled P5-8, 10,16,17. The array was developed with either biotinylated progesterone-BSA followed by HRP-linked streptavidin (left) or HRP linked sheep anti-human κ (right). The array shows that the 4 clones after antigen selection were positive for antigen binding whereas those before selection were negative, a result confirmed independently by *E. coli* expression. The anti-κ detection showed that the V_H/K fragments were expressed and immobilised for all clones.

Example 7. Coupled expression and immobilisation of luciferase on magnetic beads.

Luciferase was chosen for generation of a functional enzyme immobilised on the solid surface of magnetic beads using the PISA procedure. A construct encoding luciferase with the C-terminal double His-tag (Luci-His) was produced by PCR as described for V_H/K; as a control, luciferase DNA lacking the His-tag domain (Luci) was also PCR generated. Primers for PCR generation of luciferase DNA were as follows:

T7 Luci/back: GCA GC TAA TAC GAC TCA CTA TAG GAA CAG ACC ACC ATG *GAA* GAC GCC *AAA* AAC [SEQ ID No. 9]. The T7 promoter and Kozak signal are underlined. Italics indicate the sequence complementary to the 5' region of luciferase. ATG is the initiation codon.

Luci-linker/for: GCC ACC GCC TCT AGA GCG CAA TTT GGA CTT TCC GCC [SEQ ID No. 10]. The underlined sequence overlaps a *linker-tag/back* primer (see above) used for generation of the double (His)₆-tag domain.

After cell free expression in the presence of Ni-NTA coated magnetic beads, the latter were separated from the translation mixture and washed. Luciferase activity free in the translation mixture supernatant and immobilised on the beads was measured using a luminometer (Figure 11). While the luciferase construct lacking the His-tag domain only produced activity in the translation mixture, Luci-His generated activity both in the mixture and on the beads, demonstrating the successful immobilisation of functional luciferase through the double-His tag domain.

Example 8. Generation of protein *in situ* array (PISA) elements from immobilised DNA.

In this example, PCR DNA was immobilised on magnetic beads and used as template to generate His-tagged proteins which became immobilised *in situ* on the surface of Ni-NTA coupled wells or magnetic agarose beads.

PCR DNA fragments encoding the human anti-progesterone V_H/K fragment P5-17 or luciferase was biotinylated using a 3' primer labelled with biotin. Coupling to beads was achieved by mixing the biotinylated PCR fragment with streptavidin-linked magnetic beads (Promega) and incubating at room temperature for 30 min with gentle rotation. The DNA-coupled beads were collected and washed three times with 0.5ml phosphate buffered saline (PBS).

To create an array element of P5-17 V_H/K by PISA on a Ni-NTA coated well surface, 25 μ l of TNT mixture containing 0.02mM methionine and 0.5mM Mg acetate was mixed with P5-17 DNA-coupled beads (above) and the mixture added into the well of a Ni-NTA coated plate (Ni-NTA HisSorb strips, Qiagen). After incubation for 2 hrs with shaking, the plate was washed 100 μ l wash buffer (50mM NaH₂PO₄, 300mM NaCl, 20mM imidazole, pH 8.0), followed by a final wash with 100 μ l phosphate-buffered saline (PBS). The wells were probed with either biotinylated progesterone-BSA or HRP-linked anti-human κ chain. TNT mixture without PCR-coated beads was used as control. Figure 12(a) shows while the negative control (TNT mixture without DNA-coupled beads) had no detected binding activity, wells incubated with P5-17 DNA-coupled beads generated positive signals using either the antigen, biotinylated progesterone-BSA, and HRP-linked streptavidin, or HRP-linked anti-human κ chain.

To create an array element of luciferase by PISA on the surface of Ni-NTA-coated agarose beads, 25 μ l of TNT mixture containing 0.02mM methionine and 0.5mM Mg acetate was added to a mixture of luciferase DNA-coupled beads and Ni-NTA coated beads (Qiagen). The mixture was incubated at 30°C for 2hr with gentle shaking. The beads were washed as above and luciferase activity measured by luminometry. A TNT mixture containing luciferase DNA-coated beads, but without Ni-NTA coated beads, was used as control. Figure 12(b) shows that both supernatant and beads from the TNT reaction mixture containing DNA-coupled and Ni-NTA coated beads generated luciferase activity, whereas with the control mixture lacking Ni-NTA beads, enzyme activity was only detected in the supernatant.

These experiments demonstrate the potential to generate immobilised protein from immobilised DNA by *in vitro* synthesis (PISA method). This could be used to convert a DNA array into a protein array.

Example 9. Time-course for creating a protein *in situ* array (PISA) on Ni-NTA wells and magnetic beads.

To measure the optimal time for immobilisation of *in vitro* synthesised protein onto Ni-NTA coated wells, 100 μ l of TNT mixture containing P5-17 V_H/K DNA was subdivided into 25 μ l aliquots and added to four Ni-NTA coated wells. The mixtures were incubated at 30°C for 1, 2, 4 and 7 hours respectively. After washes, HRP-linked anti-human κ was used to detect the immobilised P5-17 antibody fragment (Figure 13a). Two hours incubation produced the highest level of V_H/K immobilisation.

To measure the optimal time for immobilisation of *in vitro* synthesised protein onto Ni-NTA linked magnetic beads, 100 μ l of TNT mixture containing luciferase DNA was subdivided into 25 μ l aliquots and added to Ni-NTA coated beads. The mixtures were incubated at 30°C for 1, 2, 4 and 7 hours. After washes, luciferase activity was measured by luminometry (Figure 13b). This revealed that while both free and immobilised luciferase activity reached a peak in activity in 2hr, free luciferase declined significantly after 2 hours.

Example 10. Immobilisation of His-tagged P5-17 V_H/K after cell free synthesis and transfer to Ni-NTA coated wells with serial dilution.

The His-tagged P5-17 V_H/K fragment was expressed in 50 μ l rabbit reticulocyte lysate. After serial two-fold dilution, 25 μ l was added in duplicate to individual Ni-NTA coated wells. After 2 h incubation at 30°C, wells were washed and P5-17 immobilisation was detected using HRP-linked anti-human κ chain (Figure 14) . The result shows that V_H/K was detectable on the well surface after up to 4 fold dilution.

Example 11. Analysis of protein immobilisation onto Ni-NTA coated beads by Western blotting.

Western blotting was applied to estimate the efficiency of protein immobilisation onto Ni-NTA coated beads. Both P5-17 human anti-progesterone V_H/K and luciferase were immobilised onto

-20-

Ni-NTA coated beads using the PISA procedure. The beads were then boiled in SDS sample buffer before running on PAGE, in parallel with supernatant protein remaining in the TNT mixture after removal of beads. After Western blotting, transferred protein was detected with HRP-linked anti-(His)₆ antibody and quantified by densitometry. The results show that 40-50% of the translated protein was eluted from Ni-NTA coated beads (Figure 15, arrowed) (cf. similar result using Ni-NTA coated wells, example 5).

Example 12. Re-use of luciferase-immobilised beads after storage.

To test if arrayed protein immobilised by the PISA procedure can be re-assayed, luciferase-immobilised beads were generated as in example 7 and luciferase activity from both free in the supernatant and immobilised on the beads was analysed in parallel. After the first measurement, the beads were washed three times with washing buffer and twice with PBS; they were resuspended in 50µl PBS and stored at -20°C for one week. The supernatant was also stored at -20°C for the same period. The second and third measurements of luciferase activity including washing and storage of the samples were performed as for the first. Figure 16 shows that PISA-immobilised luciferase could be reassayed twice with positive results.

Example 13. Efficiency of consensus sequence for both prokaryotic and eukaryotic expression.

The designed PISA construct includes a novel consensus sequence for protein translation initiation in both prokaryotic and eukaryotic systems (Figure 2). In order to validate its effectiveness, a luciferase construct containing the consensus initiation sequence was tested for protein expression in both coupled rabbit reticulocyte lysate and coupled *E. coli* S30 systems. Figure 17 shows that the sequence allowed production of functional luciferase in both systems. A comparison also showed that it generated luciferase as efficiently as the regular sequence for eukaryotic systems (Figure 17a).

E: eukaryotic initiation sequence alone.

PE: combined prokaryotic and eukaryotic initiation sequences (novel)

Example 14. Demonstration of protein-protein interactions by combination of ribosome display and PISA: Interaction of Grb2 with Vav/N-SH3.

Interaction of the signalling proteins Vav and Grb2 (Ye and Baltimore (1994) PNAS 91:12629-12633) was chosen as a model with which to demonstrate protein-protein interaction through a combination of ribosome display and PISA protein arrays. In intact mammalian haematopoietic cells, Vav binds to the adaptor molecule Grb2 to initiate signal transduction via Ras activation. Vav contains Src homology 2 and 3 domains (SH2, SH3), in the order (N-SH3)-SH2-(C-SH3). Interaction with Grb2 was first identified by yeast two-hybrid screen and subsequently shown by filter binding assay to involve a highly specific binding between the N-SH3 domain of Vav and the C-SH3 domain of Grb2 (Ye and Baltimore, 1994). The Grb2 C-SH3 domain bound the N-SH3 domain of Vav, but not the Vav/SV2-(C-SH3) domains. Ye and Baltimore also prepared GST-Grb2 coupled to glutathione-conjugated agarose beads in order to detect interaction with Vav and showed that full-length Grb2 on beads precipitated full length Vav from cell lysates.

We generated DNA for the Vav/N-SH3 fragment (comprising the N-terminal SH3 domain), Vav/SV2-C-SH3 (comprising the SH2 domain and linked C-terminal SH3 domain), and full length Grb2 by PCR, from plasmids provided by Dr Martin Turner (Babraham Institute), using primers based on their corresponding DNA sequence. In the constructs for ribosome display, a human C κ domain was assembled at the C-terminus of the protein, while in the constructs for protein array (PISA) format the double (His)₆ domain was added to the C-terminus (no C κ domain).

(a) Interaction of Vav/N-SH3 ribosome display complex with PISA-generated Grb2

Ribosome display for Vav/N-SH3 and Vav/SV2-C-SH3 proteins was by the eukaryotic method (He and Taussig (1997) Nucl. Acids Res. 25:5132). PISA-immobilised, full length Grb2 was generated on Ni-NTA coated beads by cell free synthesis as described herein. Grb2 beads were mixed with ribosome displayed Vav/N-SH3, Vav/SV2-C-SH3 or a 1:1 mixture of Vav/N-SH3 : Vav/SV2-C-SH3, respectively. The interactions were carried out on ice for 2 hr and the ribosome complexes interacting with Grb2 beads were isolated by a magnetic concentrator. After

-22-

washing the beads three times with 50µl washing buffer (PBS, 0.1% BSA, 0.05% Tween, 5mM Mg acetate) followed by two washes with water, the beads were subjected to RT-PCR using universal primers for both Vav/N-SH3 and Vav/S2-C-SH3 constructs (T7A1 and evol 4). Figure 18 shows that while DNA encoding Vav/S2-C-SH3 was not detected, Vav/N-SH3 was efficiently recovered from both Vav/N-SH3 and the 1:1 mixture, a result consistent with the reported interaction between Grb2 and Vav/N-SH3.

(b) Interaction of Grb2 ribosome display complex with PISA-generated Vav/N-SH3

To confirm the interaction observed above, the displayed format of the proteins was reversed, i.e Vav/N-SH3 and Vav/S2-C-SH3 were expressed as PISA-immobilised proteins on magnetic beads and Grb2 was in the form of a ribosome display complex. The conditions for interaction and washing were as above. Figure 19 shows that RT-PCR recovery of Grb2 DNA was much stronger from Vav/N-SH3 beads than from Vav/S2-C-SH3 beads. This again confirms a specific interaction between Grb2 and Vav/N-SH3.

(c) Selection of interacting molecules from a ribosome display library

To test if interacting molecules could be selected from a ribosome display library, mixtures of ribosome complexes comprising Vav/N-SH3 and Vav/S2-C-SH3 in 1:1, 1:2 and 1:5 ratios were produced and reacted with PISA-generated Grb2 beads. Conditions for interaction followed by washes were as in (a) above. DNA encoding the selected interacting molecules was recovered by RT-PCR. As a comparison, the unselected ribosome display libraries were also subjected to RT-PCR using the same primers as in DNA recovery. Figure 20 shows while both Vav/N-SH3 and Vav/S2-C-SH3 were amplified proportionally by RT-PCR from the original libraries, only Vav/N-SH3 was demonstrably recovered after interaction of library mixtures of different ratios with Grb2-linked beads, demonstrating selection of the interacting molecule Vav/N-SH3.

These experiments demonstrate that the combined use of ribosome display and PISA can detect protein-protein interactions and allows recovery and identification of DNA for the interacting

-23-

partners. It will be possible thereby to screen ribosome display libraries against PISA protein arrays for discovery of novel interactions.

FIGURE LEGENDS

Figure 1: Diagram of DNA constructs for *in vitro* expression and protein immobilisation.

Figure 2(a): Upstream sequence elements [SEQ ID No. 11] for eukaryotic expression, showing T7 promoter (italics), and Kozak (italics) sequences.

Figure 2(b): Upstream sequence elements [SEQ ID No. 12] for both prokaryotic and eukaryotic expression, showing T7 promoter (italics), Shine Delgarno (S/D) sequence (underlined) and Kozak (italics) sequences.

Figure 3: DNA and protein sequences of hexahistidine and flexible linker.

[SEQ ID Nos: Fig. 3a = SEQ ID Nos. 13 (DNA) and 14 (protein)

Fig. 3b = SEQ ID Nos. 15 (DNA) and 16 (protein)

Fig. 3c = SEQ ID Nos. 17 (DNA) and 18 (protein)

Fig. 3d = SEQ ID Nos. 19 (DNA) and 20 (protein).]

Figure 4: Diagram of DNA constructs for immobilisation prior to *in vitro* expression.

Figure 5: Construction of a functional protein array element *in situ* after transcription and translation of DNA *in vitro*

The P5-17 protein or control protein were synthesised from PCR DNA *in vitro* in 3 duplicated wells each and immobilised *in situ* via a double hexahistidine tag to Ni-NTA surface of HisSorb wells.

- 1: Binding of biotin-labelled progesterone-BSA, followed by HRP-streptavidin detection.
2. Binding of HRP-anti-human-κ antibodies.
3. Binding of HRP-streptavidin

-25-

Figure 6: Construction of a functional protein array element by transfer of protein to a secondary surface after transcription and translation of DNA *in vitro*

The P5-17 protein or control proteins with double hexahistidine tags were synthesised from PCR DNA *in vitro* in noncoated polystyrene wells and transferred to HisSorb wells coated with Ni-NTA. Specific binding was detected with biotinylated progesterone-bovine serum albumin (*P-BSA), biotinylated BSA (*BSA), or biotinylated carcinoembryonic antigen (*CEA) followed by HRP-streptavidin detection. Protein expression was detected with HRP-anti-human- κ antibodies.

Figure 7: Protein *in situ* array (PISA) of human single chain anti-progesterone V_H/K (P5-17). (For explanation see Results, Example 4): y=yellow colour; b=blue colour.

Figure 8: Protein *in situ* array (PISA): Western blot quantitation of expression and immobilization of human V_H/K antibody fragment. Key: C = TNT Control, T = Total, U = Unbound fraction; P = PISA bound fraction (eluted).

Figure 9: Protein *in situ* array (PISA) arrays of cloned human V_H/K fragments from a transgenic library before and after selection in ARM display: y=yellow colour; b=blue colour.

Figure 10: Illustration of construction of DNA suitable for *in vitro* protein synthesis for PISA, using single chain antibody V_H/K fragments as example.

Figure 11: *Functional assay of free and PISA-immobilised luciferase by luminometry.* Supernatant: TNT mixture after incubation with Ni-NTA magnetic agarose beads. NC: TNT mixture control lacking PCR DNA. Luci: TNT mixture containing the PCR construct encoding luciferase without double-His tag domain. Luci-His: TNT mixture containing the PCR construct encoding luciferase fused with double-His tag domain.

Figure 12: Generation of protein array elements using immobilised PCR DNA .

-26-

P5-17 human anti-progesterone V_H/K fragment was immobilised on a Ni-NTA coated well (imm P5-17) after TNT cell free synthesis from P5-17 DNA-coupled beads. Control: DNA-coupled beads omitted.

Luciferase was immobilised on Ni-NTA beads after TNT cell free synthesis from luciferase DNA-coupled beads (TNT + DNA-beads + Ni-beads). Controls: Ni-NTA beads alone; TNT + DNA-coupled beads only.

Figure 13: Time-course of cell free protein synthesis by TNT and immobilisation *in situ* on Ni-NTA coated wells and magnetic beads.

Immobilisation of P5-17 human anti-progesterone V_H/K fragment onto Ni-NTA coated wells; assay by binding of biotinylated progesterone-BSA and streptavidin-HRP.

Immobilisation of luciferase on Ni-NTA coated beads; activity measured by luminometry.

Figure 14: P5-17 V_H/K immobilisation after cell free synthesis, serial dilution and transfer to Ni-NTA coated wells. V_H/K was detected by HRP linked anti- κ antibody.

Figure 15: Western blotting analysis of protein immobilised *in situ* onto Ni-NTA coated beads after cell free transcription / translation (PISA method).

S: free protein in supernatant after removal of beads; B: protein eluted from beads.

- (a) P5-17 V_H/K
- (b) Luciferase

Figure 16: Repeated assay of PISA-immobilised luciferase on beads.

PISA-immobilised luciferase on Ni-NTA coated magnetic beads was assayed by luminometry; the beads were washed after each use and stored at -20°C for one week between 1st and 2nd, and 2nd and 3rd assays. Supernatant was stored similarly.

Figure 17: Efficiency of consensus translation initiation sequence for both prokaryotic and eukaryotic expression.

- (a) Expression in Coupled rabbit reticulocyte system
- (b) Expression in Coupled *E.coli* S30 system

-27-

Luciferase expression was assayed by luminometry.

NC: cell free TNT mixture without DNA

E: DNA construct containing eukaryotic sequence alone

PE: DNA construct containing consensus sequence

Figure 18: Interaction of Vav/N-SH3 ribosome display complex with PISA-immobilised Grb2.

The marked tracks are

NC: Solution control for RT-PCR.

1. Vav/N-SH3 ribosome complexes interacted with Grb2 beads
2. Vav/SI2-C-SH3 ribosome complexes interacted with Grb2 beads
3. Ribosome complexes, expressed from a 1:1 mixture of Vav/N-SH3 and Vav/SI2-C-SH3 DNA, interacted with Grb2 beads

Figure 19: Interaction of ribosome display complex Grb2 with PISA-immobilised Vav/N-SH3.

The marked tracks are

1. Grb2 ribosome complexes interacted with Vav/SI2-C-SH3 beads
2. Grb2 ribosome complexes interacted with Vav/N-SH3 beads

Figure 20: Selection of protein-protein interaction from a library mixture.

The ribosome complex libraries were mixtures of Vav/N-SH3 and Vav/SI2-C-SH3 in the ratios 1:1, 1:2 and 1:5 as marked.

Gel (A): RT-PCR result after interaction of the mixtures with Grb2 beads

Gel (B): RT-PCR of starting mixtures

CLAIMS

1. A protein array created by transcription and translation of DNA, or by translation of mRNA, by a cell free system *in vitro* to produce individual proteins, domains or peptides, which are distributed in a gridded format.
2. A protein array created by transcription and translation of DNA, or by translation of mRNA, by a cell free system *in vitro* to produce individual proteins, domains or peptides, which contain amino acid sequences enabling covalent or noncovalent attachment to a surface or bead, such that the proteins can be arranged in a gridded format after interaction with appropriate ligands or reagents on the surface.
3. A protein array as in claim 2 in which the proteins, domains or peptides include single or double hexahistidine sequences for immobilisation through interaction with metal ions such as nickel.
4. A protein array as in claim 2 in which the proteins, domains or peptides include specific peptide sequences or other modifications for immobilisation through recognition by antibodies or other reagents interacting with the peptide sequence or the modification.
5. A protein array as in claims 2-4, in which the DNA molecules are distributed for transcription and translation, or the mRNA molecules are distributed for translation, in a gridded format on a surface or in the presence of a bead which carries a ligand or reagent for immobilisation of the proteins, such that individual proteins become attached to the surface or bead as they are produced.
6. A protein array created, as in claim 2, by the transcription and translation of DNA molecules distributed and covalently or noncovalently bound to a surface or bead in a gridded format, where the surface or bead also carries a ligand or reagent for immobilisation of the proteins, as in claims 3 and 4, such that individual proteins become attached to the surface or bead as they are produced.

7. A protein array created, as in claim 2, by the translation of mRNA molecules distributed and covalently or noncovalently bound to a surface or bead in a gridded format, where the surface or bead also carries a ligand or reagent for immobilisation of the proteins, as in claims 3 and 4, such that individual proteins become attached to the surface or bead as they are produced.
8. A protein array as in preceding claims in which the DNA is produced by PCR from genomic DNA, cloned DNA fragments, cDNA, or other DNA template source, or by RT-PCR from mRNA
9. A protein array as in preceding claims in which the DNA for transcription and translation is cloned plasmid DNA.
10. A protein array created from DNA constructs which can be expressed in either or both prokaryotic and eukaryotic cell free systems as a result of inclusion of appropriate primer sequences.
11. A protein array as in preceding claims using a cell free system which enables coupled transcription and translation of DNA, or which allows separate transcription of DNA and translation of mRNA, and is of rabbit reticulocyte, wheat germ, yeast, bacterial or other eukaryotic or prokaryotic origin.
12. A protein array as in preceding claims in which the transcription of DNA and translation of mRNA are performed as separate reactions.
13. A protein array as in all preceding claims in which the coupled transcription/translation of DNA, or the separate translation of mRNA, is carried out in microwells or on the surface of glass, nitrocellulose or other membrane, or other supporting medium or in the presence of beads.
14. A protein array created by *in vitro* transcription and translation of DNA, or the separate translation of mRNA, as in preceding claims, in which the proteins produced are transferred in a gridded format to another surface or bead where they are bound or maintained in solution.

-30-

15. Use of a protein array created as in preceding claims to identify interactions of the arrayed proteins with other molecules, including antibodies, other proteins or domains, peptides, small ligands, cell extracts, nucleic acids, etc.
16. Use of a protein array created as in preceding claims to identify interactions of the arrayed proteins with other molecules displayed in a library, such as a phage display or ribosome display library.
17. Use of a protein array created as in preceding claims to study cellular protein expression profiles.
18. Use of a protein array created as in preceding claims to study post-translation modifications of cellular proteins.
19. A protein array created as in preceding claims in which the proteins arrayed are single chain antibody fragments.
20. A protein array created as in preceding claims in which the proteins arrayed retain and display their native function.
21. A method for creating a protein array in which
 - (a) DNA constructs, containing sequences allowing transcription and translation by a cell free system, are prepared,
 - (b) said individual DNA constructs are transcribed and translated *in vitro* using a cell free system, and
 - (c) the proteins, domains or peptides produced are distributed in a gridded format on a surface or in solution to create said protein array.
22. A method for creating a protein array in which
 - (a) DNA constructs, containing sequences allowing transcription and translation by a cell free system, are prepared,

-31-

- (b) said DNA constructs are individually immobilised at a surface,
- (c) said individual DNA constructs are transcribed and translated *in vitro* using a cell free system, and
- (d) the proteins, domains or peptides produced are distributed in a gridded format on a surface or in solution to create said protein array.

23. A method for creating a protein array in which

- (a) DNA constructs are prepared containing sequences allowing transcription and translation by a cell free system and sequences enabling covalent or noncovalent attachment of the encoded protein, domain or peptide to a surface or bead,
- (b) said DNA constructs are distributed in a gridded format,
- (c) said DNA constructs are transcribed and translated using a cell free system,
- (d) the proteins, domains or peptides produced are transferred in a gridded format to a well, surface or bead on which they become immobilised through the specified peptide sequences to create said protein array.

24. A method for creating a protein array in which

- (a) DNA constructs are prepared containing sequences allowing transcription and translation by a cell free system and sequences enabling covalent or noncovalent attachment of the encoded protein, domain or peptide to a surface or bead,
- (b) said DNA constructs are distributed in a gridded format and immobilised either covalently or noncovalently, in wells, on a surface or a bead,
- (c) said DNA constructs are transcribed and translated using a cell free system,
- (d) the proteins, domains or peptides produced are transferred in a gridded format to a well, surface or bead on which they become immobilised through the specified peptide sequences to create said protein array.

25. A method for creating a protein array in which

- (a) DNA constructs are prepared containing sequences allowing transcription and translation by a cell free system and sequences enabling covalent or noncovalent attachment of the encoded protein, domain or peptide to a surface or bead ,

-32-

(b) said DNA constructs are distributed in a gridded format into wells, onto surfaces or in the presence of a bead, any of which carries a ligand or reagent for immobilisation of the proteins,
(c) said DNA is transcribed and translated, in the gridded format, using a cell free system, and
(d) the proteins, domains or peptides produced become attached to the well, surface or bead as they are produced to create said protein array.

26. A method for creating a protein array in which

(a) DNA constructs are prepared containing sequences allowing transcription and translation by a cell free system and sequences enabling covalent or noncovalent attachment of the encoded protein, domain or peptide to a surface or bead,
(b) said DNA is distributed in a gridded format and immobilised either covalently or noncovalently, in wells, on a surface or a bead, any of which carry a ligand or reagent for immobilisation of the proteins,
(c) said DNA is transcribed and translated, in the gridded format, using a cell free system, and
(d) the proteins, domains or peptides produced become attached to the well, the surface or bead as they are produced, to create said protein array.

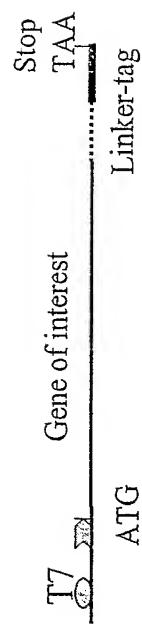
27. Methods for creating a protein array as in claims 21-26 in which the DNA is transcribed into mRNA and the mRNA translated in separate reactions.

28. A method for creating a protein array as in any of claims 21-26 in which the DNA is produced as in any of claims 8, 9 or 10.

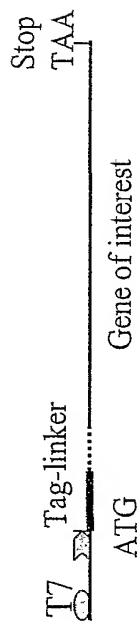
29. A method for creating a protein array as in any of claims 21-26 in which the cell free system is as in claim 11.

Figure 1

Construct with downstream (C-terminal) tag



Construct with upstream (N-terminal) tag



T7: T7 promoter; \bowtie Protein expression signal; ATG: start codon
—: tag sequence; : peptide linker

Figure 2(a) [SEQ ID No. 11]

T7 promoter
5'-GCA GCT AAT *ACG ACT CAC TAT AGG AAC AGA CCACC ATG-3'*
Kozak

Figure 2(b) [SEQ ID No. 12]

T7 promoter
5'-GCA GCT *AAT ACG ACT CAC TAT AGG GA* GAAGGAGACCACC ATG-3'
S/D Kozak

Figure 3

a. (Single hexahistidine) – linker upstream: *CAT CAC CAT CAC CAT CAC GGC GGT GGC TCT GGT GGC GGT TCT GGC GGT GGC ACC GGT GGC GGT TCT GGC GGT GGC* [SEQ ID No. 13],
encoding **HHHHHHGGGSGGSGGGTGGGSAGG** [SEQ ID No. 14].

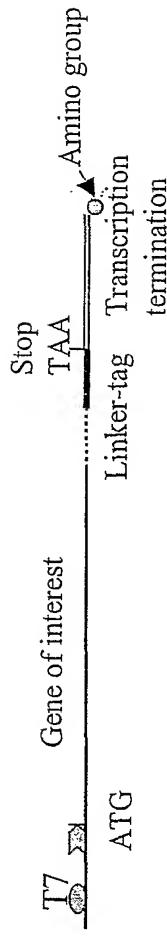
b. (Double hexahistidine) – linker upstream: *CAT CAC CAT CAC CAT CAC TCT AGA GCT TGG CGT CAC CCG CAG TTC GGT GGT CAC CAC CAC CAC CAC GGC GGT GGC TCT GGT GGC GGT TCT GGC GGT GGC ACC GGT GGC GGT TCT GGC GGT GGC* [SEQ ID No. 15],
encoding **HHHHHHHSRAWRHPQFGGHHHHHGGSGGGSGGGTGGGSAGG** [SEQ ID No. 16].

c. (Single hexahistidine) – linker downstream: *GGC GGT GGC TCT GGT GGC GGT TCT GGC GGT GGC ACC GGT GGC GGT TCT GGC GGT GGC CAT CAC CAT CAC CAT CAC* [SEQ ID No. 17],
encoding **GGGSAGGSGGGTGGGSAGGHHHHHH** [SEQ ID No. 18].

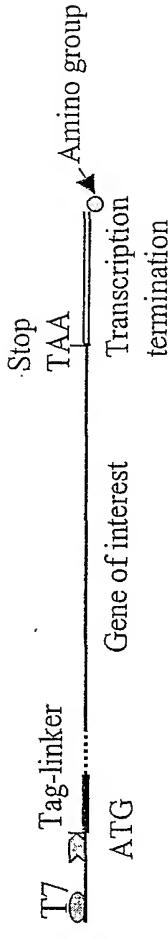
d. (Double hexahistidine) – linker downstream: *GGC GGT GGC TCT GGT GGC GGT TCT GGC GGT GGC ACC GGT GGC GGT TCT GGC GGT GGC CAT CAC CAT CAC CAT CAC TCT AGA GCT TGG CGT CAC CCG CAG TTC GGT GGT CAC CAC CAC CAC CAC* [SEQ ID No. 19],
encoding **GGGSAGGSGGGTGGGSAGGHHHHHHHSRAWRHPQFGGHHHHHHH** [SEQ ID No. 20].

Figure 4

Construct with downstream (C-terminal) tag



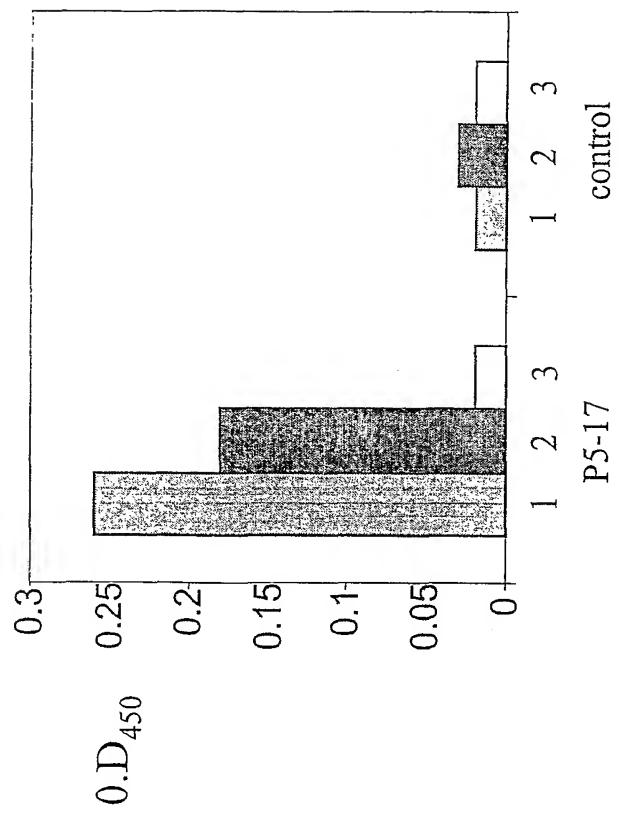
Construct with upstream (N-terminal) tag



T7: T7 promoter; \Rightarrow Protein expression signal; ATG: start codon

— : tag sequence; : peptide linker

Figure 5



6/20

Figure 6

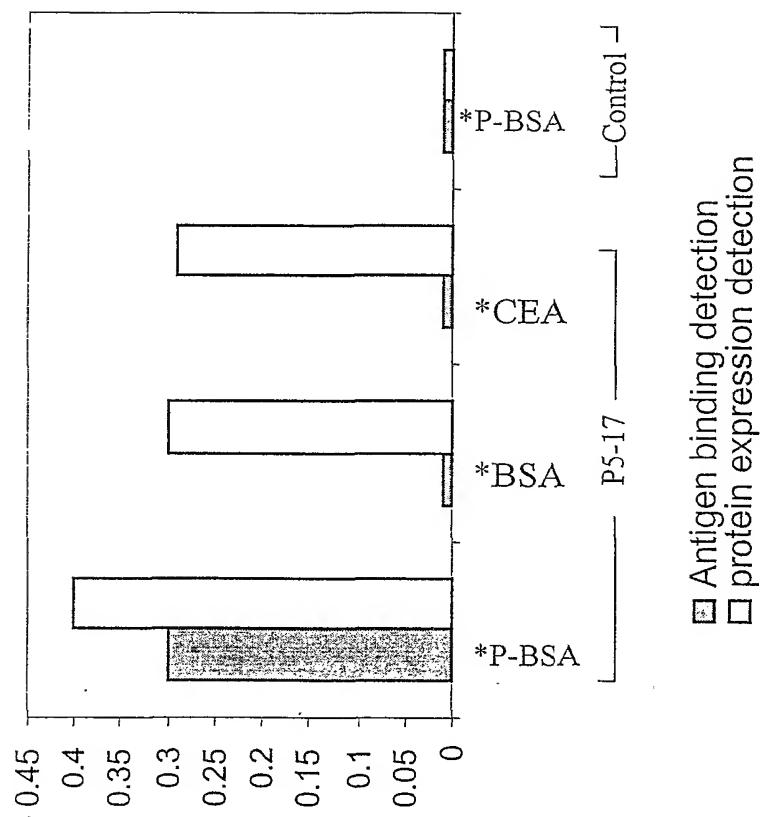


Figure 7

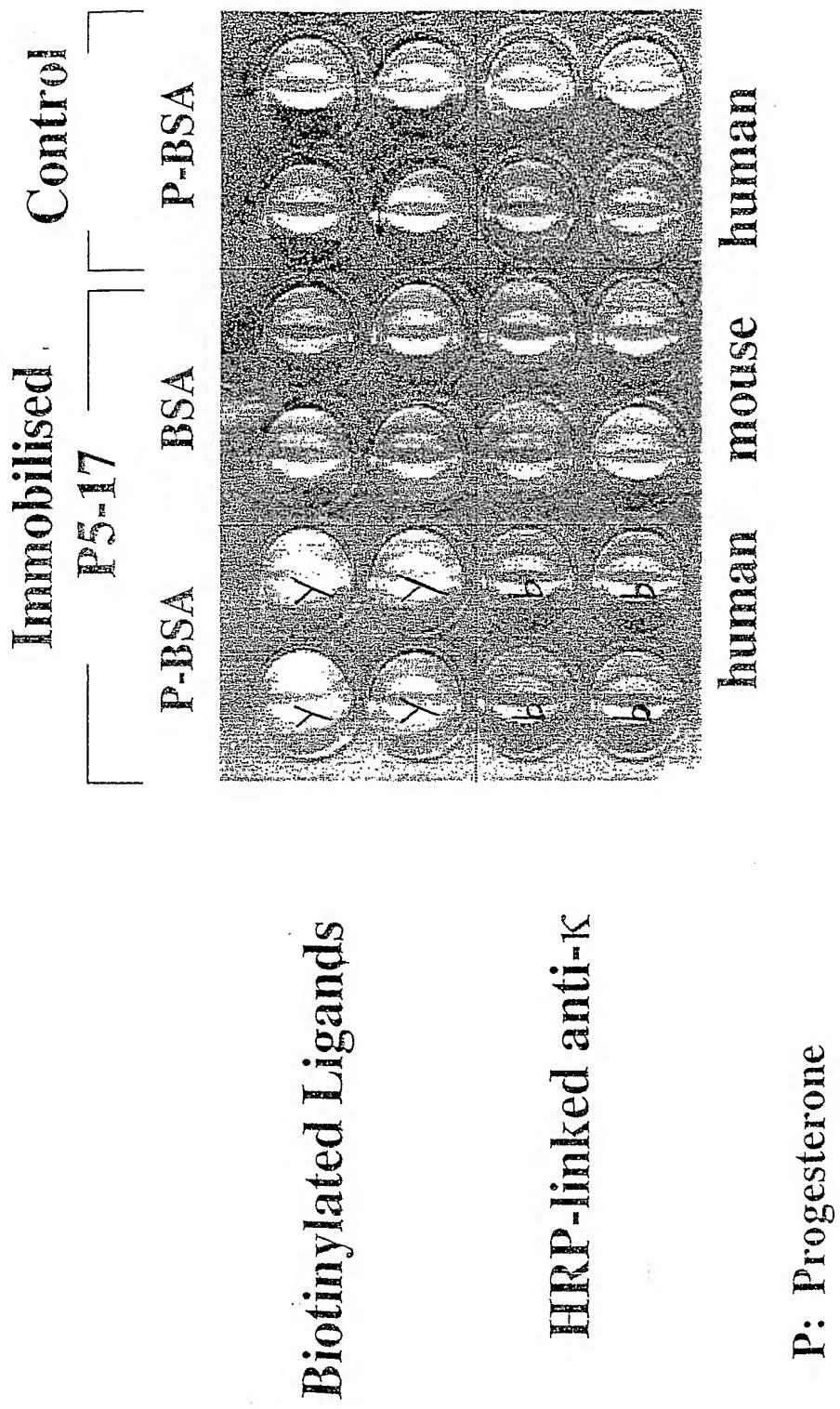


Figure 8

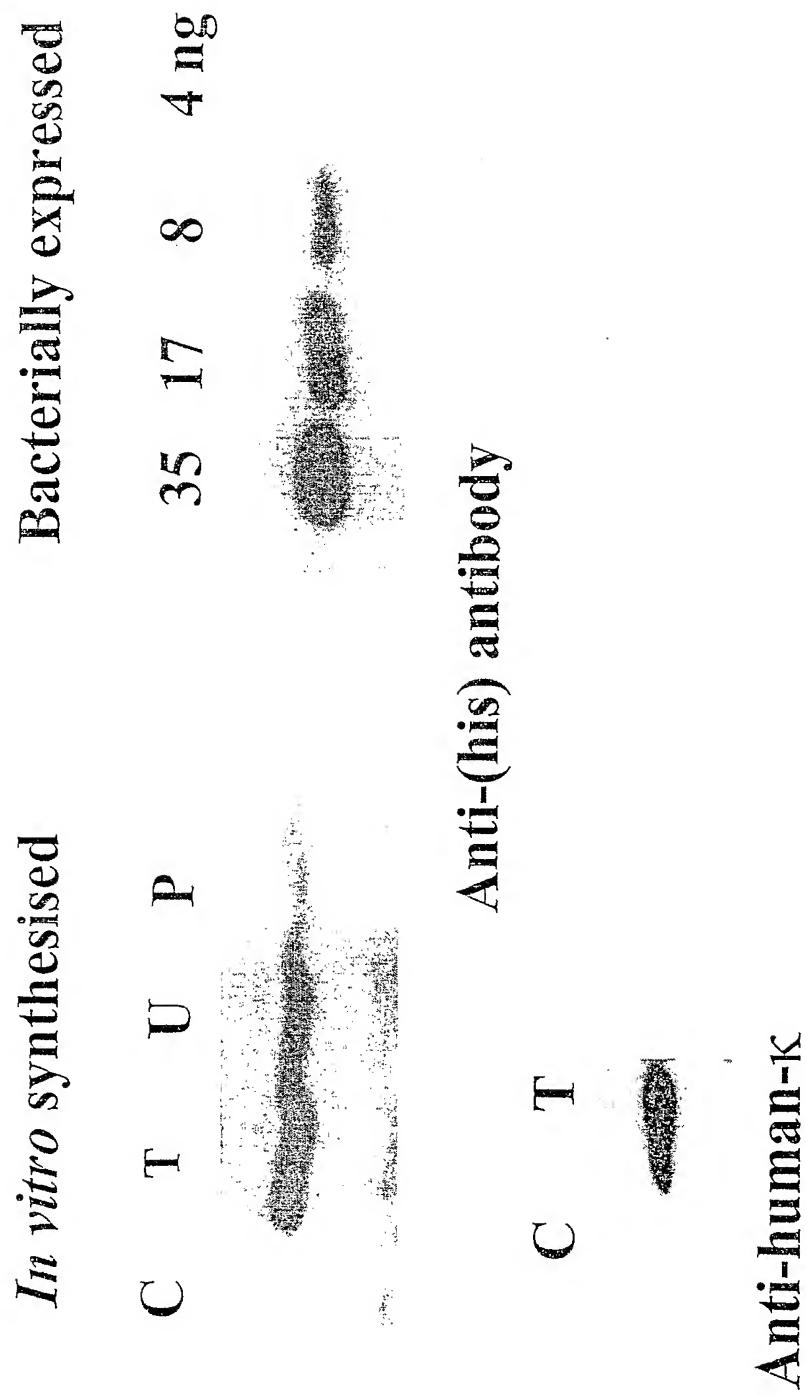


Figure 9

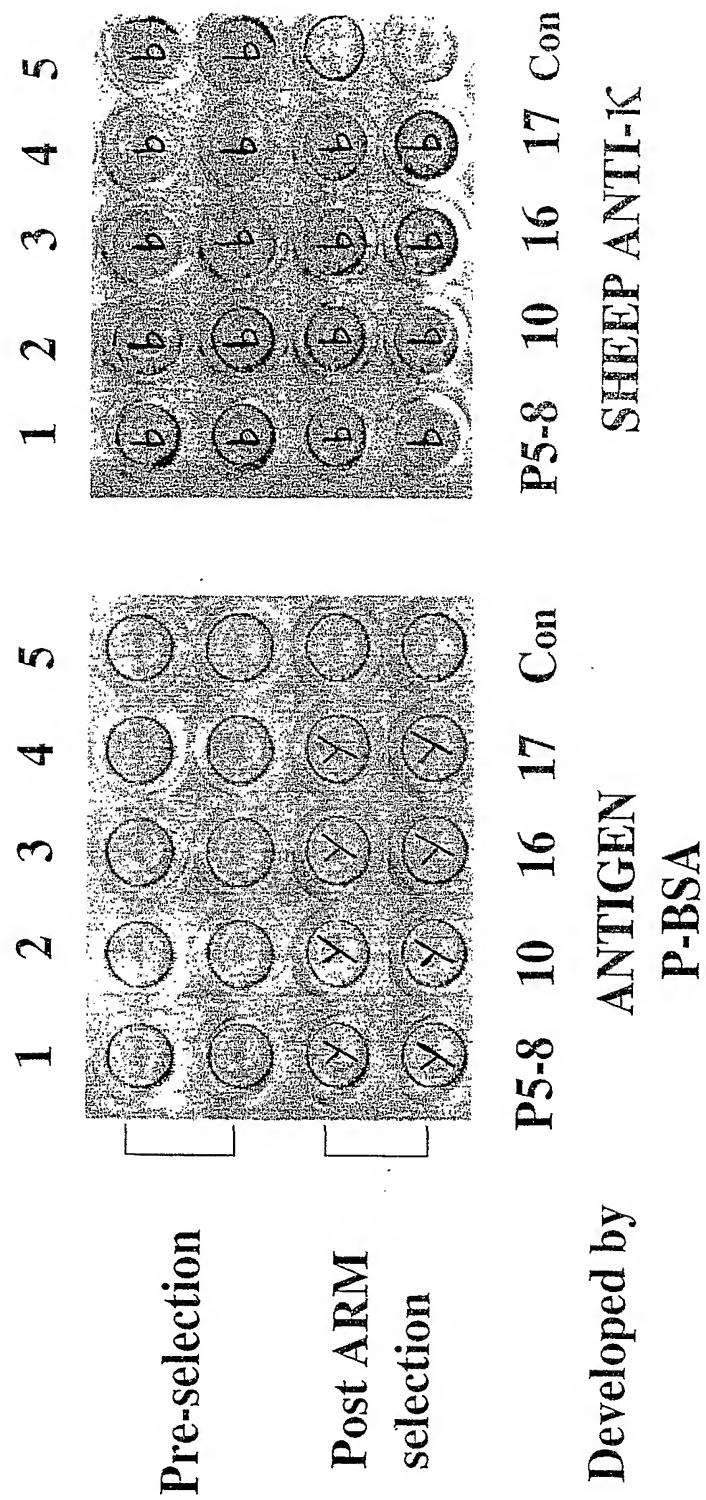


Figure 10

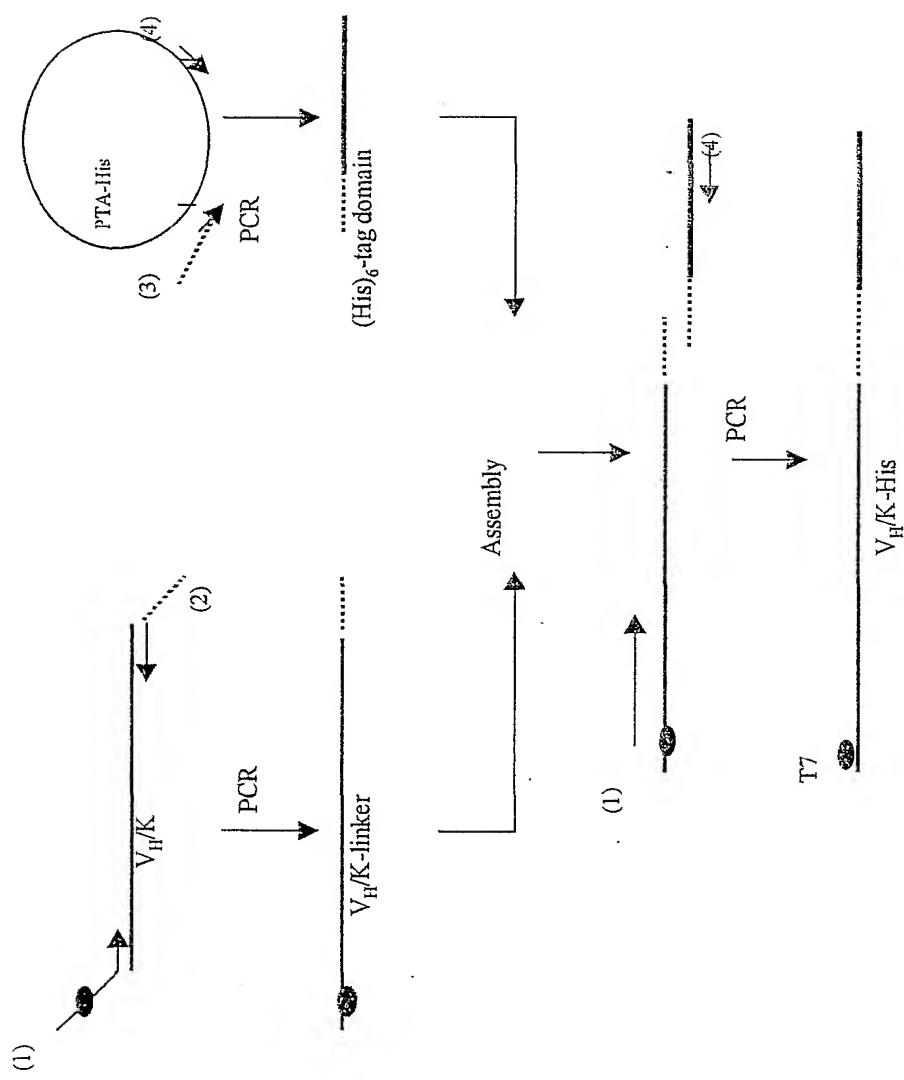


Figure 11

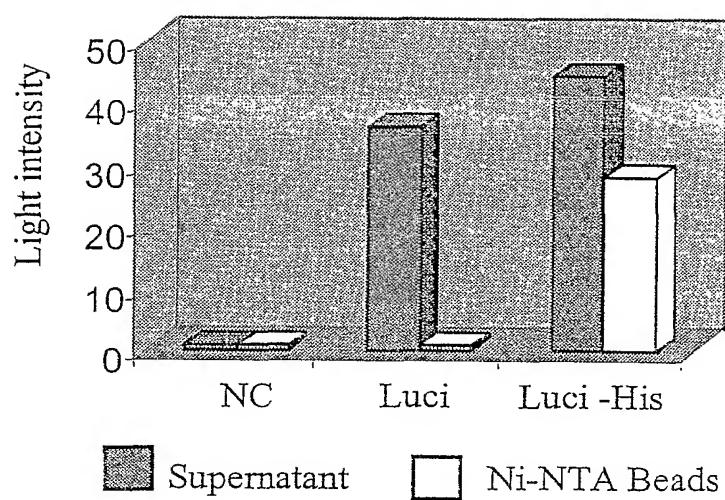


Figure 12

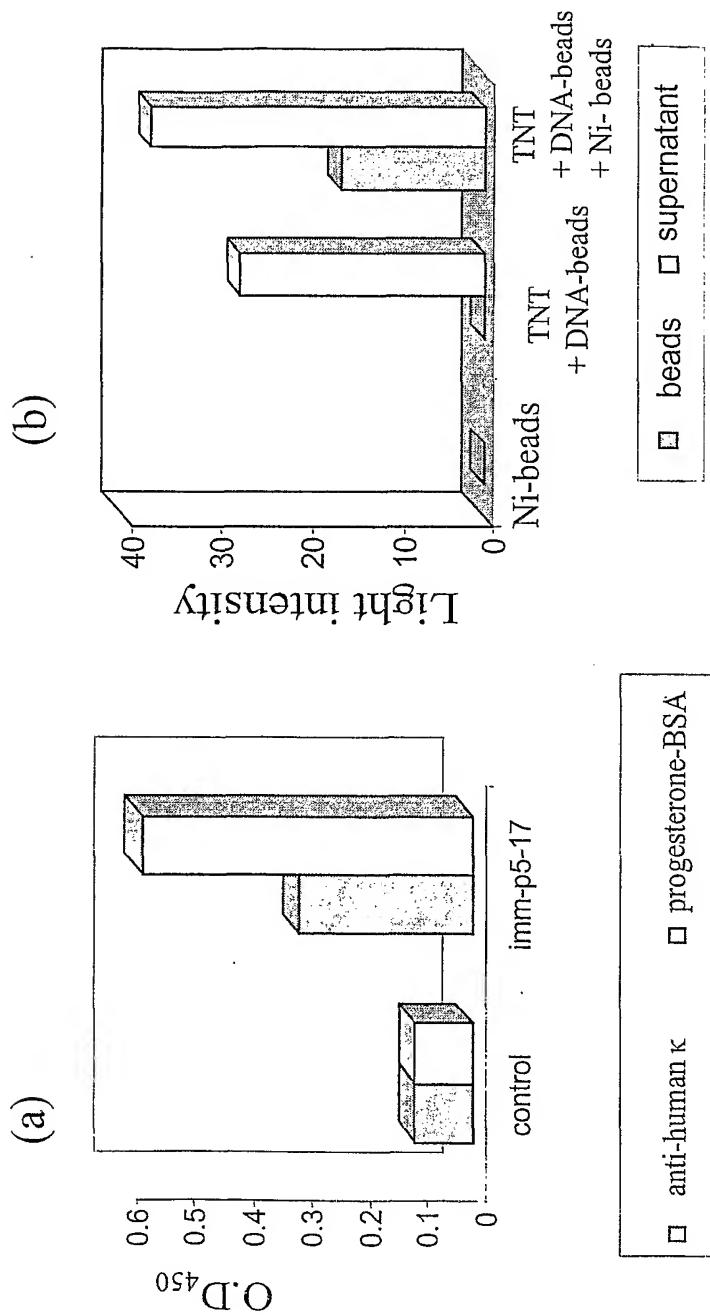


Figure 13

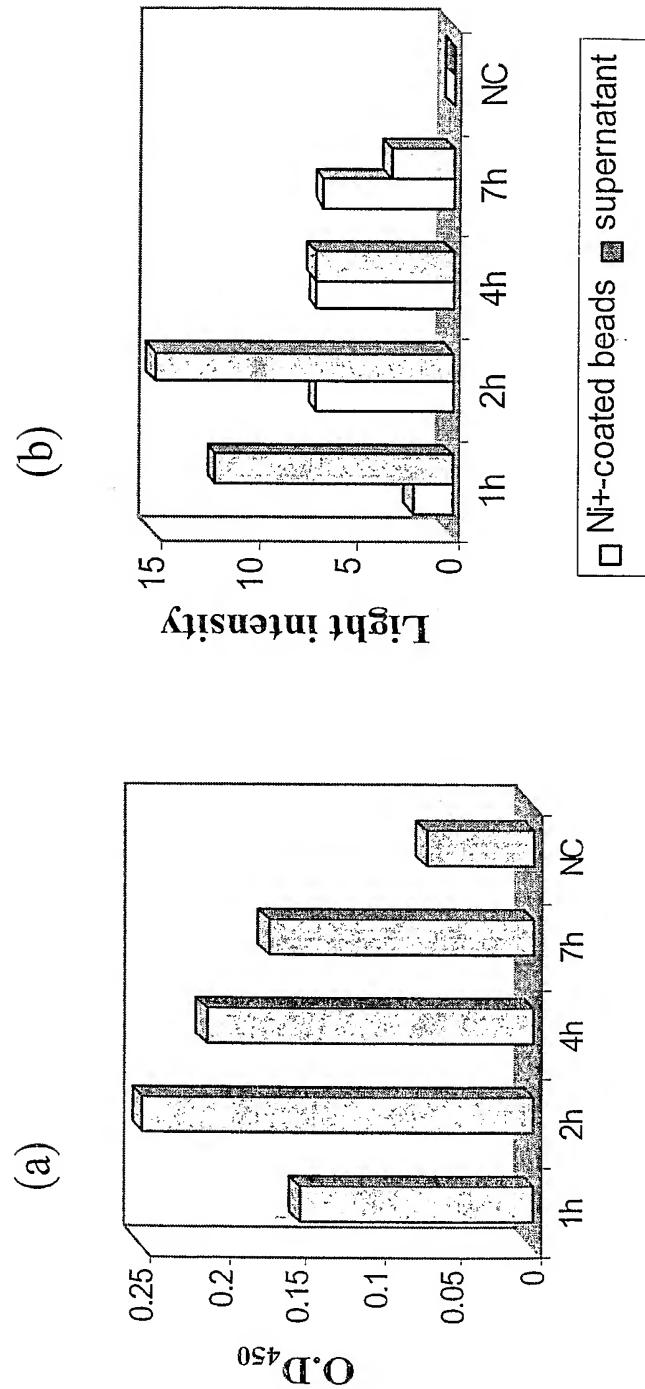


Figure 14

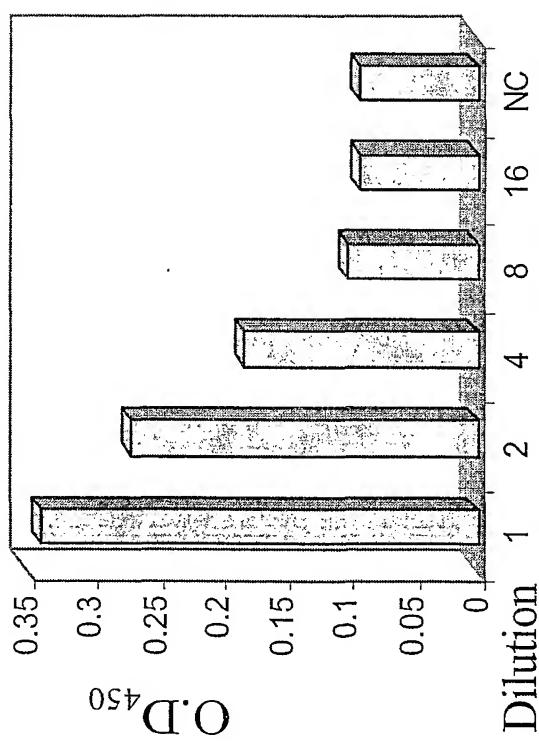


Figure 15

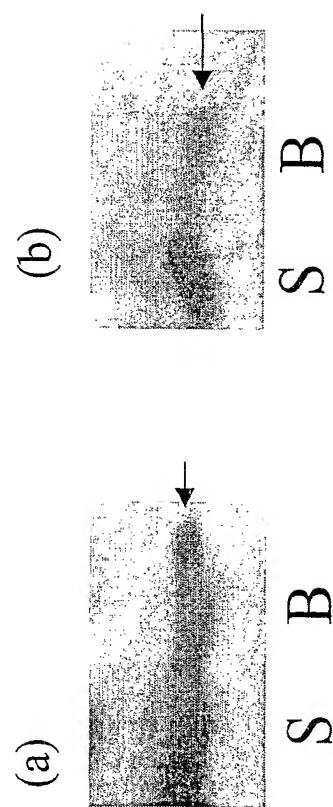


Figure 16

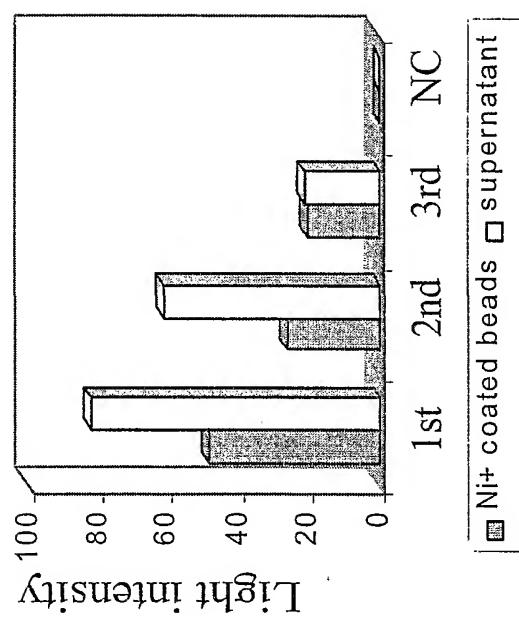


Figure 17

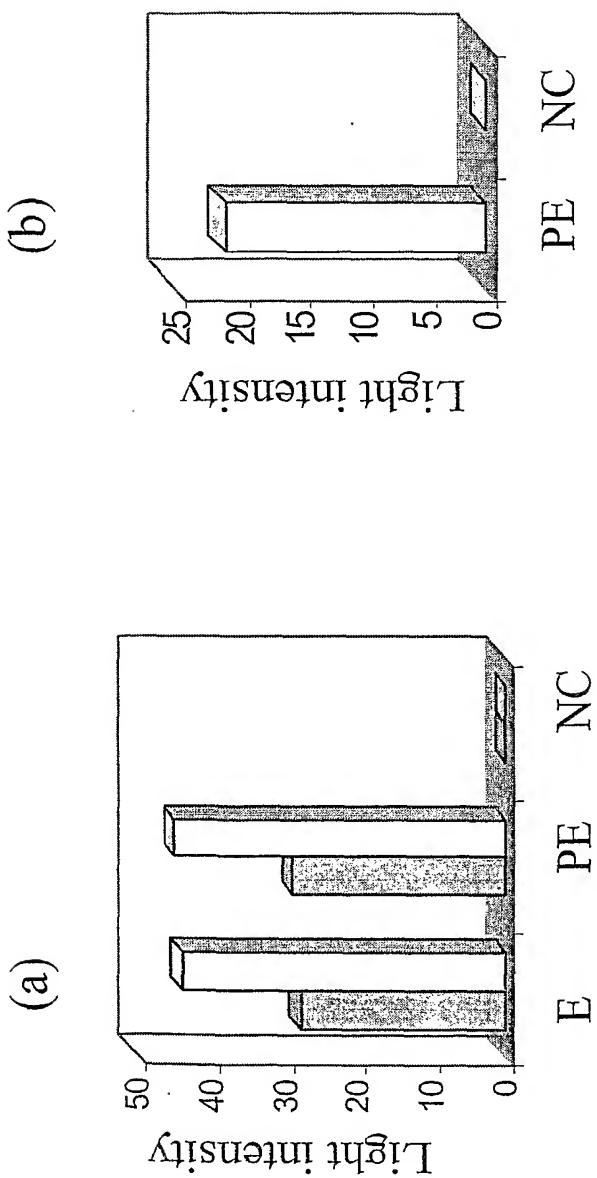


Figure 18

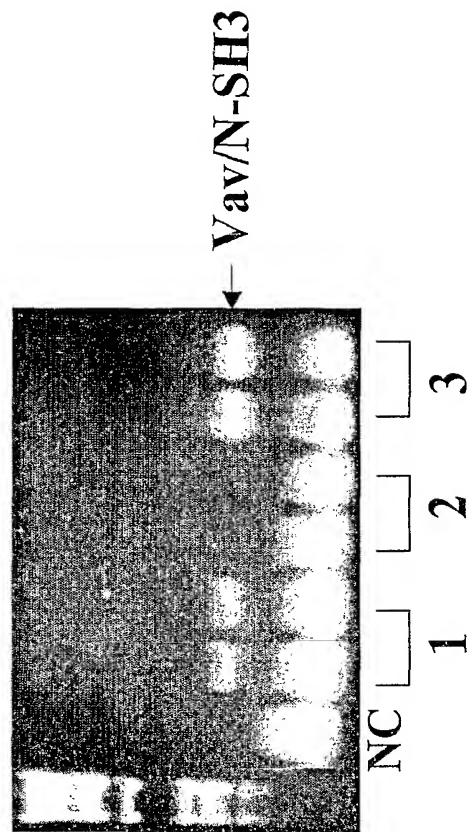


Figure 19

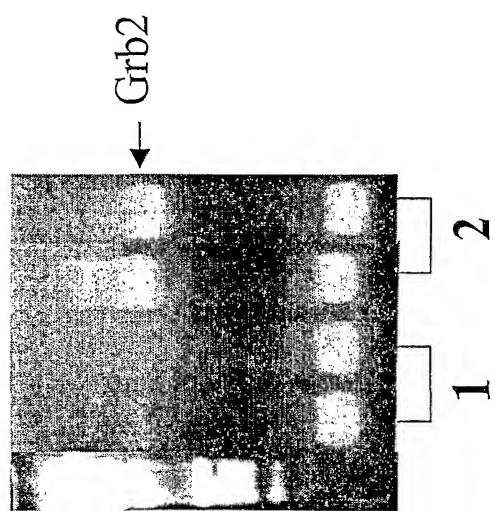
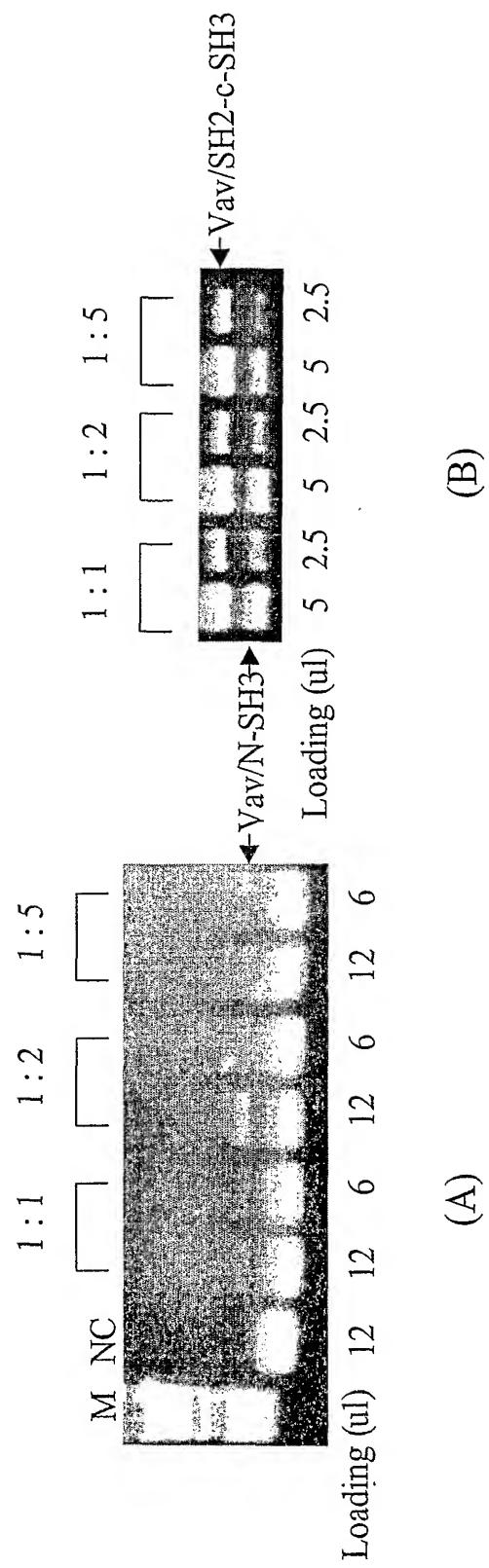


Figure 20

SEQUENCE LISTING

<110> Discerna Limited

<120> Functional Protein Arrays

<130> 42312pct/jmd/mar

<140>

<141>

<160> 20

<170> PatentIn Ver. 2.1

<210> 1

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Polypeptide

<400> 1

His His His His His His
1 5

<210> 2

<211> 23

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Polypeptide

<400> 2

His His His His His Ser Arg Ala Trp Arg His Pro Gln Phe Gly
1 5 10 15

Gly His His His His His His
20

<210> 3

<211> 6

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
Oligonucleotide

<400> 3

taataa

6

<210> 4

<211> 61

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR Primer

<400> 4

gcagctaata cgactcacta taggaacaga ccaccatgsa ggtscasctc gagsagtctg 60
g 61

<210> 5

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR Primer

<400> 5

gccaccgcct ctagagcggc tcagcgtcag ggtgctgct 39

<210> 6

<211> 83

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR Primer

<400> 6

gctctagagg cgggtggctct ggtggcggtt ctggcggtgg caccgggtggc ggttctggcg 60
gtggcaaacg ggctgatgct gca 83

<210> 7

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR Primer

<400> 7

tccggatata gttcctcc 18

<210> 8

<211> 272

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
Oligonucleotide

<400> 8

gctctagagg cgggtggctct ggtggcggtt ctggcggtgg caccgggtggc ggttctggcg 60
gtggcaaacg ggctgatgct gcacatcacc atcaccatca ctctagagct tggcgtcacc 120
cgcagttcgg tggtcaccac caccaccacc actaataaaa aaaaaaaaaaaa aaaaaaaaaaaa 180
aaaaaaaccgc tgagcaataa ctgcataaac cccttggggc ctctaaacgg gtcttgaggg 240
gtttttgct gaaaggagga actatatccg ga 272

<210> 9
<211> 53
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 9
gcagctaata cgactcacta taggaacaga ccaccatgga agacgcacaa aac 53

<210> 10
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 10
gccaccgcct ctagagcgca atttggactt tccgcc 36

<210> 11
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 11
gcagctaata cgactcacta taggaacaga ccaccatg 38

<210> 12
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 12
gcagctaata cgactcacta tagggagaag gagaccacca tg 42

<210> 13
<211> 75
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 13

catcaccatc accatcacgg cggtggctct ggtggcggtt ctggcggtgg caccgggtggc 60
gtttctggcg gtggc 75

<210> 14
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligopeptide

<400> 14
His His His His His Gly Gly Gly Ser Gly Gly Ser Gly Gly
1 5 10 15
Gly Thr Gly Gly Gly Ser Gly Gly
20 25

<210> 15
<211> 126
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 15
catcaccatc accatcaactc tagagcttgg cgtcacccgc agttcggtgg tcaccaccac 60
caccaccacg gcgggtggctc tggtggcggt tctggcggtg gcaccgggtgg cggttctggc 120
ggtggc 126

<210> 16
<211> 42
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligopeptide

<400> 16
His His His His His Ser Arg Ala Trp Arg His Pro Gln Phe Gly
1 5 10 15
Gly His His His His Gly Gly Ser Gly Gly Ser Gly
20 25 30
Gly Gly Thr Gly Gly Ser Gly Gly
35 40

<210> 17
<211> 75
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 17
ggcggtggct ctgggtggcggt ttctggcggt ggcaccggtg gcggttctgg cggtgtggccat 60
caccatcacc atcac 75

<210> 18
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligopeptide

<400> 18
Gly Gly Gly Ser Gly Gly Ser Gly Gly Gly Thr Gly Gly Ser
1 5 10 15
Gly Gly Gly His His His His His His
20 25

<210> 19
<211> 126
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 19
ggcggtggct ctgggtggcggt ttctggcggt ggcaccggtg gcggttctgg cggtgtggccat 60
caccatcacc atcactctag agcttggcggt caccgcagt tcggtgtca ccaccaccac 120
caccac 126

<210> 20
<211> 42
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligopeptide

<400> 20
Gly Gly Gly Ser Gly Gly Ser Gly Gly Gly Thr Gly Gly Ser
1 5 10 15
Gly Gly Gly His His His His His Ser Arg Ala Trp Arg His Pro .
20 25 30
Gln Phe Gly Gly His His His His His His His
35 40